

Chiedi PDF (17 giu 2024, 12:11:05)

****UPDF AI:****

****Titolo****

L'umanità sopravviverà all'IA? (Will Humanity Survive AI?)

****Autori****

Garrison Lovely

****Affiliazione****

Scrittore freelance che vive a Brooklyn

****Parole chiave****

Artificial Intelligence, AI Risk, Extinction, Deep Learning, General Artificial Intelligence, AI Safety

****Urls****

None

****Sommario****

- Il contesto di ricerca di questo articolo riguarda le preoccupazioni esistenziali legate allo sviluppo dell'Intelligenza Artificiale (IA) avanzata e il potenziale rischio di estinzione per l'umanità;
- I metodi passati includevano lo sviluppo di sistemi di IA senza una comprensione completa del loro funzionamento e senza piani di sicurezza adeguati. Questi approcci sono problematici perché potrebbero portare alla perdita di controllo umano sui sistemi autonomi di IA;
- La metodologia proposta nell'articolo si concentra sull'importanza di regolamentare e controllare lo sviluppo dell'IA per evitare rischi esistenziali. Viene suggerito di bloccare iniziative tese a costruire sistemi di IA che potrebbero soppiantare l'umanità o essere basati su interessi personalistici;
- Le prestazioni dei metodi attuali mostrano un miglioramento continuo delle capacità dell'IA, ma anche un aumento delle preoccupazioni sulla sicurezza. Nonostante i progressi, la mancanza di un piano di sicurezza solido rende difficile sostenere che le prestazioni attuali possano supportare gli obiettivi di sicurezza.

****Metodi****

- Analisi delle opinioni di esperti come Larry Page, Richard Sutton, Yoshua Bengio, Geoffrey Hinton e altri sul rischio esistenziale dell'IA;
- Discussione sulle implicazioni etiche e sociali dello sviluppo dell'IA avanzata;

- Esame delle iniziative di regolamentazione e delle lettere aperte firmate da ricercatori e personalità influenti nel campo dell'IA;
- Valutazione delle preoccupazioni immediate e future legate all'uso malevolo dell'IA e alle armi autonome letali.

****Conclusione:****

- L'importanza di questo lavoro risiede nella sua capacità di evidenziare i rischi esistenziali associati allo sviluppo incontrollato dell'IA e di promuovere un dibattito sulla necessità di regolamentazioni più rigorose;
- Punto di innovazione: l'articolo mette in luce la frammentazione delle opinioni tra esperti e la necessità di un consenso globale sulla regolamentazione dell'IA;
- Prestazione: l'articolo è efficace nel presentare una panoramica completa delle preoccupazioni esistenziali legate all'IA, ma potrebbe beneficiare di ulteriori dati empirici;
- Carico di lavoro: il lavoro richiede un'analisi approfondita delle opinioni degli esperti e delle iniziative di regolamentazione, rendendolo complesso ma necessario per comprendere appieno i rischi dell'IA.

****Domande che potrebbero interessarti:****

- Quali sono le principali preoccupazioni di Yoshua Bengio riguardo allo sviluppo dell'IA?
- Come ha reagito l'opinione pubblica americana al lancio di ChatGPT secondo i sondaggi del 2023?
- Quali sono le implicazioni etiche e sociali discusse nell'articolo riguardo all'uso delle armi autonome letali?

****UPDF AI:****

****L'evoluzione dell'Intelligenza Artificiale****

****Key Points:****

- Larry Page e Richard Sutton esprimono opinioni contrastanti sull'impatto dell'IA sull'umanità.
- Yoshua Bengio critica l'approccio di Page e Sutton, definendolo criminale.
- Preoccupazioni riguardo alla possibilità che l'IA possa portare all'estinzione dell'umanità.
- Alcuni esperti di IA come Yann LeCun e Geoffrey Hinton si uniscono alla comunità dell'x-risk per discutere dei rischi esistenziali legati all'IA.

****La visione degli esperti sull'IA****

- Larry Page e Richard Sutton hanno opinioni diverse sull'evoluzione dell'IA e sul suo impatto sull'umanità.
- Yoshua Bengio critica l'approccio di Page e Sutton, definendolo criminale.
- Esperti come Yann LeCun e Geoffrey Hinton si preoccupano dei

rischi esistenziali legati all'IA.

****UPDF AI:****

****Rischio di estinzione dovuto all'Intelligenza Artificiale****

- Il comitato etico di Google DeepMind non ha tenuto una seconda riunione dopo il 2015.
- Timnit Gebru è stata costretta a lasciare Google alla fine del 2020 a causa di una controversia su un documento sull'apprendimento automatico.
- Nel 2022, gli investimenti per la sicurezza dell'IA hanno raggiunto i 144 milioni di dollari, quattro volte di più rispetto al 2020.
- Esiste una crescente preoccupazione riguardo ai rischi a lungo termine legati all'IA, con un aumento degli investimenti nel settore.

****Percezione del rischio di estinzione nell'Intelligenza Artificiale****

- Secondo Bengio, la percezione del rischio di estinzione dovuto all'IA è aumentata nel tempo, passando dallo 0,1% al 5%.
- David Chalmers propone una tesi probabilistica, affermando che anche un rischio del 5% dovrebbe essere sufficiente per preoccuparsi.
- Esistono opinioni contrastanti sulla reale minaccia di estinzione dovuta all'IA, con alcuni scettici e altri che sottolineano i progressi significativi nel campo.
- La comunità della sicurezza dell'IA invita a "rallentare" lo sviluppo tecnologico per consentire una regolamentazione adeguata e una riflessione sulle implicazioni future.

****Cooperazione internazionale per la regolamentazione dell'IA****

- L'obiezione principale alla regolamentazione dell'IA è spesso rappresentata dalla Cina, ma esistono opinioni divergenti sulla necessità di regolamentare l'IA a livello globale.
- Alcuni esperti sostengono che la Cina si stia muovendo rapidamente nella regolamentazione dell'IA e che una cooperazione internazionale potrebbe sensibilizzare sui rischi connessi.
- Esistono proposte radicali, come mettere sotto supervisione internazionale tutti i centri dati dell'IA per evitare abusi o utilizzi dannosi.
- Biden e Xi Jinping hanno discusso della sicurezza dell'IA e del divieto di utilizzarla nei sistemi d'arma, evidenziando una crescente attenzione dei governi verso la regolamentazione dell'IA.

****Impatto sociale ed etico dell'Intelligenza Artificiale****

- Molti studiosi e attivisti si stanno mobilitando contro le armi autonome letali (LAW) e chiedono un bando internazionale.
- La sinistra politica sta affrontando sfide nel formulare posizioni chiare sull'IA del futuro, con alcune correnti che propongono di smettere di produrla per evitare rischi.
- L'AI Now Institute sottolinea che non c'è nulla di inevitabile

nell'IA e che la società deve decidere quali valori guidino lo sviluppo tecnologico.

– La creazione di un'IA di livello umano solleva importanti questioni filosofiche e politiche sulla direzione della società e sui valori da perseguire.